Attachment L

Assessment of Measurement Errors and Spatial Uncertainties Associated with Bathymetric Change Analysis

1.0. Introduction

Sediment stability evaluation is a standard part of remedy selection at large contaminated sediment superfund sites. These evaluations often combine the outputs of deterministic sediment transport models, empirical mass balance arguments, geochemical evaluations such as core dating, and morphological assessment through bathymetric change analysis. Each of these techniques has its own sets of strengths and weaknesses, but each technique should include uncertainty analysis that can be incorporated directly into remedy selection.

At the Lower Passaic River, bathymetric surveys were conducted in 1989, 1995, 1996, 1997, 1999, 2001, 2002, 2004, 2007 and 2008. These bathymetric surveys provide a unique opportunity for detailed temporal comparison of bathymetric changes over both long and short temporal intervals, as well as at varying spatial scales. Comparison of these survey data requires consideration of the uncertainties inherent in the survey technologies, the level of precision in each survey, and the uncertainties in spatial interpolation. Uncertainty refers to a quantity or process associated with the measurement or analysis that is not known. Therefore, the direct analysis of bathymetry data presents unique statistical issues that must be carefully worked through in order to develop reliable results and interpretation. This attachment discusses the measurement errors inherent in bathymetric surveys, and describes a geostatistical approach, called conditional simulation, that was used to quantify uncertainty bounds in the estimates of elevation and volumetric changes. This technique allows reliable identification of erosion and accretion areas and determination of sediment budget estimates in the Lower Passaic River.

2.0. Measurements Errors and Their Influence on Bathymetric Change Analysis

2.1. General Classification of Measurements Errors

The old carpenters' adage says that it is best to "measure twice and cut once." This is much in alignment with how statisticians view the world. The "measure twice" idea recognizes that there are small errors in any measurement and the precision/reproducibility of results can be improved by making repeated measurements. If, on the other hand, the carpenter unknowingly cut off the end of his ruler, any number of measurements cannot compensate for the fact that each measurement is systematically off or biased due to the shortened ruler. A measurement with low bias is considered accurate. In general, the overall error in any measurement is due to components of both bias and precision and these can be defined as follows:

- Precision/reproducibility represents the error between repeated measurements without regard to the true value. This may be called repeatability.

- Bias represents the systematic difference between measured and true values. Measurements with low bias are termed accurate.

2.2. Consideration of Bias and Precision/Reproducibility in Statistical Analyses

A primary objective of statistical estimation is to provide estimates of unknown population parameters and to characterize the precision with which these unknown parameters are estimated from sample measurements. Standard statistical analyses generally do not expressly incorporate corrections for systematic bias or inaccurate measurements.

In general, bias may affect comparisons of estimates among measurements like bathymetric surveys, but typically does not adversely influence precision estimates. When measurements are susceptible to systematic bias or inaccuracy, independent calibration studies are used to minimize those biases. Alternatively, measurements at survey line crossings provide a means to directly assess the extent to which survey data may be biased. Small errors in measurements at survey line crossings indicate that biases have been effectively minimized through field calibration.

2.3. Error types that influence Bathymetric Change Analysis in the Lower Passaic River

Modern bathymetric surveys require coordination of a host of measurement devices including sonar systems, satellite global positioning system (GPS) receivers, tide gages, and devices that measure a boat's orientation in the water, among others. Accuracy and precision of water depth measurements require synchronization of measurements in time and space, calibration to known depth targets, careful attention to field conditions and constant monitoring and reevaluation of processes so that calibrations are maintained throughout each survey period. In a memorandum for record dated March 2, 2010 called "Survey Accuracies on the Passaic River," US Army Corps of Engineers (USACE) Engineering Research and Development Center (ERDC) reviewers identified 9 error sources thought to be of primary importance to evaluating the bathymetric survey data at the Passaic. These nine errors are listed in Table 1 and they are categorized relative to their potential to affect bias and reproducibility. Apparently questions raised by the reviewers were asked of Ocean Surveys Inc. (OSI) and Gahagan and Bryant Associates (GBA), the contractors responsible for some of the surveys. According to their responses (OSI, 2010; GBA, 2010) it appears that the listed sources of error that were expected to cause bias (*i.e.*, vertical offsets) were known to the contractors and efforts were made to minimize these sources of error. However, the reviewers did not request information regarding all 9 sources of error, so it is not possible to fully determine the standards to which OSI and GBA surveys were conducted. However, it appears from their responses that they were fully aware of the types of errors known to affect bathymetric surveys and that for issues they discussed in their responses, they appeared to exceed surveying quality standards outlined by the reviewers.

In 2007, a single beam survey consisting of traverse and longitudinal lines was conducted in the Lower Passaic River, providing basis for evaluating point differences at the survey line crossings. In addition, this single beam survey was immediately followed by a multibeam survey in the river, providing further opportunity to evaluate uncertainty in the precision of point measurements through the difference between the single beam soundings and corresponding multibeam soundings at corresponding horizontal locations. Analysis of these measurement uncertainties is presented in Section 5.8 below. (An additional multibeam survey was conducted in 2008, but this work has been evaluated only preliminarily as discussed in Section 6.0).

In addition to these errors that influence individual depth measurements, bathymetric survey comparisons are also complicated by the fact that survey lines are not typically aligned, making direct subtraction of individual measurements difficult. If, in addition, the survey extents do not match well, apparent differences in survey elevations may be confounded with differences in distributions among non-overlapping survey areas. Data acquisition procedures need to be designed to minimize the potential for unintentional biases to influence results. Statistical analyses intended to compare bathymetric surveys need to properly account for uncertainties due to random measurement errors and lack of spatial alignment of survey lines.

3.0. Approaches for Determining Bathymetric Change

Because survey lines among different surveys are typically not precisely aligned it is common for analysts to first interpolate survey data and then to subtract interpolated surfaces. These interpolated surfaces are constructed by estimation of the bathymetric surface at a regular grid of locations using weighted averages of the depth measurements. A variety of methods are used for interpolation, including inverse distance weighting, triangular irregular networks or kriging. When data are relatively dense, such as for bathymetric surveys, each method provides similar estimates of the surface elevation, whereas only the geostatistical method, kriging, provides a mechanism to quantify the precision of interpolated values. The estimation variance provides an estimate of the precision of interpolated values in the same way that a sample variance quantifies precision of the sample mean. The estimation variance is estimated from equations derived for this purpose.

There are two factors limiting the utility of the estimation variance derived from kriging:

1. the variance is a function of sampling density, independent of local variability of measurements, and
2. the sampling distribution of the estimated values is unknown in all but the rare situation where measurements are normally distributed.

Because of these limitations, more satisfying methods are needed to develop inferential statements about changes in bathymetry.

In more conventional circumstances when sample data are non-normally distributed, precluding conventional statistical methods, a re-sampling technique known as “bootstrapping” is often used to develop inferences to the mean or other functions of the data. Because bathymetry measurements are not statistically independent, these bootstrapping procedures are not applicable. This lack of normality and dependence of observations is a common characteristic in mining and petroleum engineering settings, so practitioners in those disciplines developed a technique known as conditional simulation for re-sampling dependent data. Conditional simulation provides a robust alternative by replacing mathematical derivation with the power of replicate sampling from probability distributions that are consistent with sample data.

4.0. Overview of Conditional Simulation

Conditional simulation is a re-sampling procedure for spatially dependent data, providing a robust tool for quantifying precision of functions of spatial data. The method is that is well-established in the geostatistical and statistical literature, as well as a variety of techniques and specialized applications, have evolved primarily out of mining and petroleum exploration.

The primary components of a conditional simulation are: 1) the frequency distribution (histogram) of sample data, 2) a model describing spatial correlation, 3) a regression model for any large scale spatial patterns in the data, and 4) transfer functions of mapped data to which inference is desired. The conditional simulation components can be grouped into two steps: 1) parameter estimation, followed by 2) sampling from the estimated probability distribution to build distributions for output transfer functions of mapped values (see Figure 1).

Importantly, parameters necessary for the conditional simulation are calculated directly from sample data—parameters are not generally chosen or estimated from other studies as might be the case for some parameters necessary for mechanistic models. In that sense conditional simulation is an empirical analysis producing a probability model consistent with the observed data, from which equally likely maps (surfaces) are generated consistent with the estimated probability model. Note that these surfaces maintain the same variability and spatial correlation as the original data, so they are not smooth like an interpolation, but rather they wander randomly with the constraint that these irregular surfaces interpolate the known data and cannot fluctuate more erratically than indicated by the measured spatial correlation of the original data.

4.1. Example Illustrations of Conditional Simulation

<u>Example 1</u>

Consider two hypothetical bathymetric surveys conducted in years 1 and 2 with observations as indicated by red and blue circles in Panel A of Figure 2. Consider the problem of determining the likelihood of erosion of sediments on the intervals from 100 to 200 (Block 1) and from 200 to 300 (Block 2) as well as the full interval from 100 to 300 (Blocks 1 and 2 combined). A typical approach would be to interpolate the data and then subtract interpolated surfaces. This approach results in mean elevation differences of -0.19 feet, -0.47 feet and -0.33 feet for Blocks 1, 2 and 1 and 2 combined, respectively. However, this approach does not provide a way to quantify the precision of these estimates.

Also shown in Figure 2 are two realizations generated by conditional simulation showing that the elevations may wander between sampling locations in ways that are not captured by the interpolated surfaces. Mean change in elevation is calculated for this pair of synthetic surfaces by sampling the surface at specific locations, subtracting each pair of values and averaging the differences. However, because there are many such surfaces consistent with the statistical properties that interpolate the sample locations, it is necessary to recreate many such pairs of surfaces so that their distribution can be developed. Panel C of Figure 2 shows another pair of equally likely realizations that are used to estimate the elevation change. The estimated erosion in Block 2 based on realizations from Panel C would be much larger than the estimate based on Panel B realizations.

Figure 3 shows a set of 5 synthetic surfaces in Panel A which provides an indication of the amount of uncertainty in the interpolated lines, and finally Panel B shows the full simulated envelope of equally likely surfaces. The uncertainty bands are widest midway between, and are narrower in proximity to sample locations. Also, simulated elevations do not perfectly interpolate the sample values, in recognition that measurements are not perfectly known. In geostatistics this is referred to as “nugget effect” which is discussed further in Section 4.2 below.

The uncertainty bands are interesting in and of themselves, but it is the distribution of the estimated elevation change on each block that is of primary interest. Figure 4 shows histograms of 1000 estimated differences for each of Blocks 1, 2, and 1 and 2 combined. The plots also include a vertical line indicating 6 inches of erosion. The histogram for Block 1 is centered slightly to the left of 0, at -0.16 feet, with the sampling distribution ranging from approximately -1.1 feet to 0.9 feet. This indicates that the estimated change in elevation is small relative to the spread in its sampling distribution and that the stability of these sediments would have to be considered uncertain.

For Block 2, the distribution is centered on -0.45 feet with 5th and 95th percentiles of -0.72 and -0.06 respectively. This indicates that one could be 95percent confident that sediments in this area had eroded over the time period between the two years. Interestingly, in this situation the probability of at least 6 inches of erosion is less than 50 percent so this block would be classified as “stable” based on the approach currently under consideration for the Lower Passaic.

For Blocks 1 and 2 combined, the results are essentially intermediate to those for Blocks 1 and 2 separately.

Figure 5 summarizes the same mean differences as cumulative probability distributions providing an easy way to evaluate the probability of a range of levels of mean erosion. For example the probability of a loss of 1 foot or more of sediment can be seen to be negligible for both blocks. The probability of at least some deposition ranges from a high of 30 percent for Block 1 to approximately 95 to 98 percent for Block 1 and Blocks 1 and 2 combined. Again it should be emphasized that in spite of the low probability of deposition, Block 1 would be classified as “stable” under the current classification system under consideration for the Lower Passaic.

<u>Example 2</u>

Figure 6 shows another set of synthetic realizations that are much more variable than the original set. The probability distribution for this example was the same as that in the first example, with the exception that the variance of the simulation was increased from 0.2 square feet to 1 square feet to illustrate the behavior of more variable simulated surfaces. The amplitude of fluctuations is greater for these more variable surfaces resulting in less power to discern small differences. It can also be seen that the magnitude of estimates of erosion have wider potential values. One might argue that this could induce larger estimates of erosion; however, this is not the case because the central tendency of the surfaces is defined by the observed data, constraining the simulation mean to be unbiased to the data.

Figure 7 shows the uncertainty envelope about the interpolated line and it is wider than that for the original set of realizations. Also, the tendency of the uncertainty envelope to narrow in proximity to sample locations is less pronounced.

Figure 8 shows that the histograms for mean differences are centered similarly to the previous example with means of -0.18, -0.45 and -0.32 respectively, but the spread of the sampling distribution is much wider, reflecting the greater level of uncertainty in estimated erosion due to greater variability in the underlying processes. This greater level of uncertainty is reflected in Figure 9, which shows cumulative probability distribution for average elevation change from blocks 1, 2 and block 1 and 2 combined. Probability of some erosion (*i.e.,* negative change) is approximately 60% to 80% respectively for blocks 1, 2 and (1 and 2) combined. Yet the probability of more than 6 inches of erosion is less than 50% so these blocks would be classified as not erosional.

4.2. Parameter Estimation

Conditional simulation analysis requires estimation of a probability model that captures the substantive features and patterns in observed data. This could include a very complex conceptual model, but in general the primary features are the large-scale trends such as the general shape of the river channel, continuity of the surface being modeled and the tendency for erratic fluctuations at small scales (*i.e.*, nugget effect).

Large-scale trends are typically modeled first using a regression analysis to identify important covariates that may be predictive of bathymetry elevations. Secondarily, the residuals from this regression model are investigated for “spatial correlation” using a statistical dissimilarity measure called the semi-variogram. Figure 10 provides a schematic of the semivariogram model indicating the important aspects and the parameters that must be estimated.

There are three primary parameters of the variogram that determine the nature of simulated surfaces: the sill, range of influence and nugget effect. The sample semivariogram is a plot of variance among samples plotted as a function of distance between points. In general, when data are positively spatially autocorelated (*i.e.*, when proximate samples are more similar than distant samples) the semivariogram usually increases with distance reaching a horizontal asymptote, the sill, which signifies the variance among samples that are too distant to be correlated—the variance of spatially independent samples. The distance at which the sill is achieved is called the range of influence and it signifies the distance at which pairs of samples are statistically uncorrelated. Finally, the vertical intercept is called the nugget effect. Taken literally, the vertical intercept represents the variance of measurements taken at the same location—separated by no distance. It is the estimated nugget effect that represents the variance associated with all components of measurement error, except those that manifest themselves as global offsets. The nugget effect cannot capture inaccuracies in measurements that manifest themselves as global offsets among surveys because such biases subtract out when calculating measures of variance within an individual survey.

5.0. Application of Conditional Simulation to Bathymetric Change Analysis in Lower Passaic River

Bathymetric survey measurements in the Lower Passaic River before 2007, were integrated and reported on 10-foot intervals along transects, and transects were spaced at approximately 100-foot intervals perpendicular to river flow. Because of this sampling design, resulting data are spatially correlated across flow and relatively variable along flow. In addition, transect locations typically varied by plus or minus 18 feet among surveys. Together, these factors cause temporal changes of interest to be potentially confounded with cross- and along-flow spatial variation, as well as potential artifacts due to sample misalignment. The nature of this potential confounding is complicated by the scale of aggregation over which comparisons are made. For example comparisons of mean elevation within relatively large areas, such as the hydrodynamic model grid cells being used for the Lower Passaic, are less influenced by this lack of spatial alignment than are comparisons at smaller scales.

Bathymetric change analysis typically entails interpolation of bathymetric surfaces, followed by subtraction of resulting interpolated surfaces. This approach is expected to be relatively accurate for estimation of net deposition; however, the effects of smoothing due to the interpolation can also be expected to under-estimate the amount of material eroded at unsampled locations. This bias can adversely affect estimates of resuspension that are important to understanding temporal trends in contaminant concentration in surface sediments. Future risks and remedial selection are greatly dependent on understanding future contaminant concentrations in surface sediments.

Because bathymetry measurements are spatially correlated, because sampling locations among surveys (*e.g.*, in different years) are spatially unaligned, and because comparisons are of interest at varying spatial scales, standard statistical procedures requiring random independent samples are unsatisfying and potentially inaccurate for some comparisons of interest. As an alternative to more conventional bathymetric change analysis methods, we used the geostatistical technique known as conditional simulation (Journel and Huijbregts, 1978) to directly study the temporal change, spatial heterogeneity and spatial scale with regard to understanding sediment stability. These geostatistical techniques were selected because a probabilistic model is used to explicitly account for the effects of spatial correlation, temporal change, varying scales of interest and mismatch of sampling locations among surveys. The proposed methods are particularly powerful in that the results can be expressed as maps of the probability of sediment instability (erosion) at varying temporal and spatial scales, particularly at scales smaller than that used for the hydrodynamic and sediment transport model grid for the Lower Passaic.

As noted above, geostatistical simulation procedures were developed in the mining industry in response to the need to estimate parameters and their uncertainty distributions from data that are spatially correlated. The methods explicitly provide tools to estimate parameters that vary spatially, and to propagate uncertainty over ranges of scales of interest. For example, for calibration of hydrodynamic models, the average bathymetry is of interest at the scale of the sediment transport model domain grid cells. Conversely, areas of erosion and deposition may occur at sub-model-domain scales that if ignored could cause remedial evaluation to be biased due to the effects of smoothing the model domain. Conditional simulation provides a probability distribution at each node of a fine-mesh grid (*i.e.*, 3 foot spacing) that can then be integrated, providing error distributions for any function of the bathymetry, defined over specified areas of interest.

Conditional simulation methods have not been used widely at other superfund sites in the past, but have been recognized recently as a powerful tool facilitating analysis of uncertainty associated with metrics commonly used for remedial selection. The methods have been used widely in the mining industry, where propagating uncertainty through complicated transfer functions of geological data is important to cost analysis and reserve forecasting. As remedy selection progresses at large Superfund sediment sites with high remedial costs, the effort necessary to conduct a rigorous geostatistical simulation analysis of bathymetry and contaminant data will likely become more common. Recently, Kern *et al.,* (2009) used conditional simulation for the Fox River Superfund Site to refine dredge prism designs. Based on those analyses, it was found that additional sampling within planned dredging areas could be used to reduce the costs of removal of non-targeted sediments and from leaving targeted sediments behind. Recommendations of the simulation analysis were implemented in 2008, confirming that significant amounts of uncontaminated sediments were within the lateral and vertical boundaries of the 60 percent design dredge prisms (Barabas *et al.,* 2009). Dredge prism refinements developed from the simulation analysis and subsequent confirming field sampling are expected to net $5 to $10 million in savings in the first year of a 10-year dredging project.

In mining applications, data on the magnitude of tens to hundreds of drill holes and hundreds to low thousands of samples would be considered more than adequate to support a conditional simulation with hundreds of millions of investors' dollars at risk. In petroleum exploration, conditional simulation studies are conducted with orders of magnitude less data, again with extremely large sums of shareholders' money at risk. In contrast, for the Lower Passaic River geostatistical models (conditional simulations), parameters were calculated from 15 to 20 thousand single beam soundings for each survey and models were validated against over 2.5 million multi-beam soundings. These numbers of samples for parameter estimation and independent validation data are unheard of in any other setting.

For the Lower Passaic River, we found that, year over year, erosion and deposition were estimated to be nearly neutral in most areas when aggregation occurred over large spatial areas. Conversely, when considered at higher resolution (*i.e.*, smaller scales) we found that there were many areas with relatively high probability of erosion or deposition within larger, apparently stable, portions of the bottom equivalent to sediment transport model cells. We also used the analysis to identify areas that were erosional in at least one of the 36 possible pairs of 9 surveys taken over 18 years (1989 through 2007); and also to identify areas that were erosional at some points in time, but were later depositional.

The application of conditional simulation to bathymetry data at the Lower Passaic River required several data handling procedures, including: 1) grid straightening, 2) detrending, 3) semivariogram analysis, kriging and simulation and 4) calculating sediment erosion and deposition at scales of interest. These are described below.

5.1. Grid Straightening

Geostatistical procedures require estimation of the spatial relationships among sampling locations in order to determine how best to simulate or interpolate to unsampled locations. In typical geological and geographical settings, these relationships are measured along geographic coordinate axes and summarized by the semivariogram—a function that quantifies variance as a function of distance. For positively spatially correlated processes, variance is an increasing function of distance. In other words, dissimilarity is expected to increase with distance between sample locations. However, when studies involve data collected within riverine or estuarine systems, spatial relationships are determined largely by flow direction. In general, it is expected that sample locations along similar flow paths would be more similar in value than those that are located across flow directions. For bathymetry this can be seen by the obvious orientation of depth contours that parallel the direction of flow.

Because the Passaic River is sinuous, variogram analysis based on geographic coordinates would not properly capture the spatial relationships that are expected to be driven by flow direction. To correct this situation, the geographic coordinates were transformed to along- and cross-flow coordinates using the Schwarz–Christoffel Toolbox (Driscoll, 1996) as implemented in MATLAB© (The Math Works, 1998).

The coordinate transformation (Figure 11) was done as follows: 1) the physical boundary was defined by the two banks and lines marking the northern [River Mile (RM)7.5] and southern (RM0.5) ends of the more spatially constrained surveys; 2) curvilinear grid lines parallel and perpendicular to the flow were constructed using the Schwarz–Christoffel conformal mapping; 3) the Schwarz–Christoffel Toolbox was used to transform all data locations from rectangular (Euclidean) coordinate to the constructed curvilinear system. One unit of distance in the new curvilinear scale corresponds to an average of approximately 150 feet in the Euclidean scale. All geostatistical analyses were performed in the transformed coordinate system.

5.2. Exploratory Data Analysis and Transformations

The basic assumptions of the geostatistical model applied are that the bathymetric elevations can be mathematically transformed to a population that has:

1. a constant mean and variance across the area of interest,

2. spatial correlation that is a function of distance and direction, and

3. a normal distribution.

The first two conditions are the assumptions of second order stationarity (Cressie, 1993) and assumption three leads to the multigaussian model described by Deutsch and Journel (1998) and also by Goovaerts (1997) with applications in natural resources characterization. This assumption is similar to the assumption of constant variance in a traditional regression analysis. The purpose of the transformation and detrending of the bathymetry data described below is analogous to the variance stabilizing transformations commonly applied prior to a regression analysis.

5.2.1. Histograms

Figures 12a through 12h show the histograms and summary statistics of the single beam bathymetry data from 1989-2004 surveys. It can be seen that the distributions are somewhat left-skewed, and tend to have an over-abundance of measurements between 0 and -5 feet below sea level. Such deviation from normality is expected to require some level of transformation prior to application of the multi-Gaussian model for simulations; the normal score transformation was used in this case to correct for non-normality.

5.2.2. Detrending

For each survey, it was found that the cross-flow semivariograms have much shorter ranges of influence and also much higher maximum values. This is an example of zonal anisotropy (Journel and Huijbregts 1978; p 182) caused by directional differences in variation. It is not surprising that elevations would vary more rapidly across flow than along flow due to the quasi-U shaped profile of the river cross-sections. In geostatistical terms, this is indicative of a non-stationarity of the mean that could adversely affect inferences if not modeled carefully. In order to minimize the potential effects of strong zonal anisotropy on simulation results, the data were detrended.

Bathymetry elevations in the Lower Passaic River exhibit a strong U-shaped channel profile with a thalweg that meanders within the channel, tending to be closer to outside bends in areas of higher curvature. Elevations also tend to decline gradually with distance downstream. These large-scale trends would violate the constant mean assumption described above for the geostatistical model; thus their patterns were estimated and subtracted from the original elevations so that the resulting residuals could be treated as a second-order stationary process. Detrending is a common approach when analyzing spatial data (Cressie, 1993 p. 46) and is part of the process of partitioning data into large and small-scale fluctuations that are of interest. In this application, the primary interest is in understanding relatively small (3 to 6-inch) changes in elevation among surveys, so larger-scale spatial variation associated with channel cross-sections would tend to mask these smaller fluctuations of primary interest.

Large-scale spatial patterns in the bathymetry are well understood—including an approximately U-shaped river cross-section and a general downward sloping trend in the along-flow direction. It is also understood that the position of the minimum elevation in each cross-section (the thalweg) varies, making it difficult to model using a simple polynomial function of the geographic coordinates. Polynomial regressions were tested with little success, resulting in very low R-squared values (less than 20 percent; see Figure 12i-12p) and the residuals also show strong zonal anisotropy (see Figure 12j). As a substitute for more conventional trend surface approaches, the large-scale trend was estimated by calculating the average elevation within relatively large rectangular grid cells, and then smoothing the estimated cell means with a moving window average (Isaaks and Srivastava, 1989). An example of the resulting trend surface for the 1995 bathymetry can be seen in Figure 13. The objective of detrending was to partition the bathymetry into a large-scale trend component and a second-order stationary residual process. The size of the rectangular cells was selected by iterating on the semivariograms, starting with relatively large cells and decreasing them until the zonal anisotropy in the semivariograms was eliminated—consistent with the assumptions of constant mean and variance.

5.3. Semivariogram Analysis

Bathymetry elevations were detrended as described above and semivariograms were recalculated for each survey. The P-field simulation algorithm used for these studies requires semivariograms for both uniform and normal score transformed data (Srivastava, 1992). Therefore the detrended residuals were transformed to uniform and normal scores and semivariograms were calculated for each. The resulting empirical semivariograms and fitted positive definite models can be seen in Figures 14a-14p. It can be seen that the along- and cross-flow sills are similar for these directional variograms, indicating that the zonal anisotropy has been greatly reduced or eliminated by detrending. Semivariograms were constructed using the *gamv* procedure in Geostatistical Software Library (GSLIB) (Deutsch and Journel, 1998).

5.4. Kriging Analysis

Another intermediate step in the simulation process is development of a Kriged estimate of the mean and kriging variance for each unsampled location using the normal scores transformed residuals and corresponding semivariograms. This analysis was conducted using the kt3d procedure in GSLIB (Deutsch and Journel, 1998). The simple kriging option was selected. The estimated mean and kriging variances were used to estimate the conditional cumulative distribution at each location under the assumptions of the multi-Gaussian model. In effect the cumulative conditional distribution is assumed to be normal with means and variances given by the Kriged estimates.

5.5. Simulation Algorithm

Conditional simulation has been understood to be a potential tool for investigating uncertainty associated with remedial alternatives selection and design at large complex contaminated sediment sites, but the better known algorithms available to practitioners, such as sequential Gaussian simulation (Ripley, 1987; Deutch and Journel, 1998), tend to be extremely computer- and time-intensive, often limiting their application. In the mid 1980s to 1990s, methods were developed at the University of Wyoming (Borgman *et al.*, 1984) that reduced the necessary computational time of spatial simulation through application of the fast Fourier transform (FFT). Kern and Borgman (1997) described the algorithm in detail, demonstrated the accuracy of the method, and compared it with a sequential algorithm for reproduction of second-order moments and computational speed.

The frequency domain methods are so termed, because spatially correlated data are transformed to a series of two-dimensional sequences of sine and cosine functions (*i.e.*, their frequency components) which are statistically independent. Because the Fourier coefficients are statistically independent, the space domain simulation problem of generating many correlated variables with constant variance is reduced to simulation of a vector of independent random Fourier coefficients with frequency-dependent variances. This vector of independent random variables is then "reorganized" using the inverse FFT to produce a space-domain simulation with the specified covariance relations. Effectively, the sequential procedure is replaced by one application of the inverse FFT to the full vector.

The method is faster in very practical terms. For example, when simulating values at N spatial locations, the computational time for the more widely known Sequential Gaussian Simulation procedures (SGSIM; Deutch and Journel, 1998) is proportional to N^2, while the computational time for the FFT method is proportional N x ln(N), so for large simulation problems, such as for the bathymetry analysis where N > 1,000,000 spatial locations, the FFT method reduces computation time dramatically. A single simulated realization of 1,000,000 spatial locations requires approximately 2.6 seconds with the FFT method. A similar calculation using SGSIM or other sequential algorithms would require approximately 45 minutes per realization. In practical terms, including post-processing of simulation results, the significant reduction in computational time of the FFT approach dramatically increases the feasibility for simulation studies that can be handled within reasonable time frames. After all, conditional simulation is a Monte Carlo technique requiring many realizations to provide accurate inference. Depending on the size of the spatial domain, the FFT method can be used to generate many hundreds to thousands of realizations in minutes to hours, while the competing algorithms would require days. In this study, 500 realizations from each survey were constructed and post-processed.

One might ask how many realizations are adequate to estimate the things like the probability of erosion used in this application. A rough way to consider this question is to consider that probabilities associated with each cell in the simulation grid could be treated as proportions from a binomial distribution. If one defines *"p"* to be the estimated proportion of realizations in which erosion was observed, the standard error of the estimated *p* is given by $se(p) = \sqrt{p(1-p)/K}$, where *K* is the number of realizations. So for estimated probabilities near 0.7, the standard error for *K*=500 realizations is 0.02. This indicates that if the simulations were repeated using the same histogram and semivariogram parameters, one would expect probabilities ranging from 0.68 to 0.72.

Direct conditional simulation with the FFT method as described by Borgman (1984) is feasible for small numbers of conditioning data (*i.e.*, sampled locations) relative to the number of simulation nodes. A typical desktop or laptop computer with 3 gigabytes of memory can reasonably handle up to approximately 1000 sample points and one million simulation nodes. Bathymetry data are unique among other types of environmental data in that there are typically hundreds of thousands of sample locations for conditioning, so the direct FFT approach would not be feasible without dramatically larger capacity computer resources. To work around this problem an alternative simulation algorithm known as P-Field (Srivastava, 1992; Srivastava and Froidevaux, 2005) was selected because the algorithm can be combined with the FFT, resulting in an accurate simulation algorithm with speed and efficiency similar to the full FFT approach.

The P-Field algorithm includes estimation of the conditional cumulative distribution function (CCDF) at each of the unsampled locations conditional on the observed sample data through kriging. Once these CCDFs are constructed, each realization is sampled by selecting a uniform random value at each location and inverting the CCDF evaluated at the selected uniform value to produce a simulated value in the original scale. The idea behind the P-Field algorithm is that the uniform variables are spatially correlated to insure that the values drawn from the CCDFs reproduce the spatial correlation of the original untransformed data. Simply selecting uniform random values for each location without regard to spatial correlation would result in simulated realizations that reproduced the sample histogram, but with inflated nugget effect. So the FFT algorithm is used to produce unconditional realizations of the uniform random variables that are properly spatially correlated. This step in the process is the most computationally intensive for a sequential algorithm, requiring minutes per each of the 500 realizations, whereas the FFT method requires just 3 to 6 seconds per realization. The method is extremely fast and each realization reproduces the original sample histogram and semivariogram.

5.6. Reproduction of Data Histograms and Semivariograms

To compare bathymetric elevations among surveys, an equally likely surface was simulated for each survey of interest, and the elevations were subtracted on each pixel in the simulation grid. These differences were compared with difference cutoffs of interest, such as 3 inches, 6 inches or 12 inches, and the number of times a simulated difference exceeded each cutoff value was recorded for each pixel. These frequencies were then divided by the total number of simulations (K=500), representing the probability of erosion at these selected intervals of interest. Differences were also integrated over larger areas of interest, such as those corresponding to sediment transport model grid cells, providing probability distributions for the amount of erosion averaged over these cells. The results of these and other comparisons are provided in Chapter 11 of the Conceptual Site Model.

Figure 15 provides three realizations of the simulated elevation surface for 1995. It can be seen that the sample data constrain the range of variation, yet there are distinct differences among simulated surfaces that represent the uncertainty remaining in bathymetry interpolation.

Figure 16 illustrates the histograms of simulated and actual bathymetric data. It can be seen that the simulation algorithm reproduces the input sample histogram as desired. It should be noted that this is a comparison of the histogram of all simulated locations to the data histogram, as opposed to the subset of simulated locations at which comparisons were constructed. This latter comparison cannot be developed without biasing artifacts of sampling, because the distribution of samples is not independent of the subset of cells used for comparison, and because the comparison subset is restricted to deeper areas than those represented by the sample data. Because of this biasing one would not expect the sample histogram to match the simulated histogram within a sub-area of the river.

Figure 17 shows the comparison between the theoretical along- and cross-flow semivariogram models and actual semivariograms calculated from 20 simulated surfaces. It can be seen that the semivariogram is reasonably well reproduced for both directions. If anything, these 20 realizations may slightly under-state the along-flow variogram which would tend to result in understatement of the magnitude of long-range fluctuations, potentially creating a small degree of smoothing in the along-flow direction.

5.7. Summary of Bathymetry Change Observations – Maps of Erosion (minimum 6 and 12 inches) and Deposition in the Lower Passaic River

Conditional simulation results were used to estimate the probability of erosion within the lower Passaic River (RM0 - RM8) based on a series of comparisons of bathymetric surveys. To aid in visualizing these results, maps were produced, showing places where there is a 70 percent probability of exceeding several erosion thresholds (*e.g.*, any erosion, at least 3 inches of erosion, at least 6 inches of erosion, at least 12 inches of erosion). The selection of a 70 percent probability threshold represents a very generous criterion by identifying an area as erosional only if there are more than 7 out of 10 chances that it will be so.

The 70 percent probability threshold is higher than is typical of USEPA's approaches on other sites, and there is important precedent for use of much lower thresholds. For example:

1. 50 percent and 40 percent probability thresholds, respectively, were implemented at the Fox River (Kern et al., 2008) and Hudson River (Kern, 2005) Superfund sites to delineate PCB contamination.

2. For the Pizza Road dioxin site, a USEPA Superfund site in Missouri, Saito and Goovaerts (2002) developed a critical probability threshold of 65 percent for remediation.

In addition, Barabas *et al.*, (2001) previously estimated a critical probability threshold of 54 percent for 2,3,7,8-TCDD contamination in the Lower Passaic River at a concentration level of 10 ppt. Generally speaking, USEPA guidance (1989, pg. 2-3) argues that a Superfund Site is to be assumed not to have attained cleanup standards (*i.e.*, in this case that contaminated sediments are unstable) unless "substantial evidence" shows otherwise at a high level of confidence.

Based on these observations from other Superfund sites, USEPA guidance and the bathymetric evidence that most areas in the Lower Passaic River have been subject to meaningful erosion events, it is possible that the 70 percent probability criterion used in depicting erosion represents an approach that may not be viewed by some stakeholders as protective. Thus, the resulting identified portion of the riverbed is best viewed as representing the minimum justifiable area to be addressed. Even so, this assumption covers nearly the entire river bottom in RM0-8.

An "optimal" probability threshold can be selected based on costs associated with false positive and false negative errors in delineating contaminant distributions. Section 5.9 presents a validation analysis that determines appropriate probability thresholds for various bathymetric elevation difference criteria.

The maps for erosion show the areas with at least one occurrence of a 6-inch erosional event with a greater than or equal to 70 percent probability of occurrence (Figures 18a through e). Figures 19a through 19e show the areas subject to at least one 12-inch erosional event at a similar level of probability. Each erosion map shows two demarcated areas: 1) areas with at least 1 erosion occurrence over all 36 possible combinations of surveys performed between 1989 and 2007 (underlying layer shown in red), and 2) areas with at least 1 erosion occurrence over the eight sequential survey pairs only (1989-1995; 1996-1996; 1997-1997; 1997-1999; 1999-2001; 2001-2002; 2002-2004; 2004-2007), shown as hatched areas in green (6 inches) or blue (12 inches).

For the deposition mapping analysis, the objective was to depict the frequency of significant deposition observations and long-term sediment accumulation in areas that have not been subject to any significant erosional events in the past based on the available surveys. A significant depositional observation is defined as having at least 70 percent probability of any deposition. Figures 20a through 20e display the frequency of significant deposition in areas that have not been subject to any significant erosion in the past, when considering all 36 possible survey pairs (*i.e.*, in non-significant erosional areas). In this map, frequency of observation is not the same as the number of actual depositional events since there may be multiple observations of the same depositional event. Similarly, Figures 21a through 21e display the frequency of significant deposition in areas that have not been subject to any significant erosion in the past when only sequential survey pairs are compared. This figure can be considered to represent the minimum number of depositional events since longer term changes are not captured by these comparisons. Figure 22a through Figure 22e shows the long-term elevation changes (1989 - 2007) in areas that have not been subject to any significant erosion in the past based on the 36 possible survey combinations.

The observations from these maps can be summarized as follows:

- Significant observations of erosion of 12 inches or more occurs mostly on the outside of the curves along the river (Figures 19a through 19e). This area is about 30 percent of the total area simulated when all 36 survey pairs are considered and about 17 percent for sequential pairs.

- The footprint of at least one 6-inch erosional event is more widespread and generally extends from the channel center line to the edges on the outside of curves along the river (Figures 18a through 18e). The area under this scenario is about 54 percent of the total simulated area, when all 36 survey pairs are considered. When only sequential pairs are considered, the area subject to at least one 6-inch erosional event is about 40 percent of the total area simulated.

- The erosional areas depicted in Figures 18 and 19 likely represent minimum areas, since the bathymetric surveys are unlikely to have been frequent enough to capture all such events.

- Areas without observable erosion are generally limited to below RM5 in areas on the insides of the curves along the river (Figures 20 and 21). However, these areas are not consistently depositional at 70 percent or more probability, since the frequencies of deposition are typically less than the total number of survey comparisons. That is, these areas do not always accumulate sediment between surveys but have not been observed to erode.

- Areas without observable erosion represent only 20 percent of the area simulated.
- In areas without observable erosion, the highest long-term sediment accumulation rates [more than 4 feet over the 18 year period or greater than 2.7 inches per year (in/yr)] occurs in small patches between RM2 and RM0.
- Overall, these maps suggest that significant erosion is widespread in the river with more than half of the area having undergone at least one 6-inch sediment erosion bed change over the past 18 years. While deposition observations have occurred, no significant areas can be delineated where deposition has been consistent over time.

5.8. Accounting for Bathymetric Survey Data Errors in Conditional Simulation

Bathymetric survey measurements are potentially influenced by errors associated with both accuracy and precision. As described above, statistical procedures, including conditional simulation are specifically developed for the purpose of quantifying precision of estimates derived from sample data. Also like other statistical procedures, unrepresentative or inaccurate data cannot be resurrected through statistical analysis. Rather, when measurements are particularly susceptible to unintended biases (broad-area consistent offsets in one direction) preemptive actions are necessary in the form of frequent calibration, testing and cross checking—and it appears that OSI and GBA took these precautions based on their responses to questions posed by ERDC staff conducting an independent review of the bathymetric surveys (USACE ERDC, 2010).

5.8.1. Inaccuracy and Systematic Bias

The ERDC reviewers identified 9 potential errors that could cause a combination of bias and precision problems in acoustic survey data in general (USACE ERDC, 2010), but only requested information from OSI and GBA regarding a subset of these 9 types of errors. OSI and GBA appear to have responded adequately to all of the concerns posed to them by ERDC, indicating that they are fully aware of the kinds of errors identified by ERDC. However, because only selected questions were posed to OSI and GBA, there remains a hypothetical gap in the record regarding the potential for large area biases to be present in the survey data. It is clear that both OSI and GBA conducted frequent calibration tests and probably exceeded ERDC standards for minimizing systematic biases and must therefore be presumed to have provided accurate data.

It does not follow from their responses, however, that these data are necessarily highly precise; but then, that is the focus of the statistical analysis. The conditional simulation combines a careful partitioning of the structural large-scale features through detrending and river straightening, spatial correlation analysis and estimation of random error directly in the statistical analysis—after all the sampling distribution of erosion estimates is the object of the analysis.

Errors in positional accuracy, heave, pitch and roll all result in misstatement of the apparent elevation in a way that is expected to occur approximately randomly throughout the survey. The effect of these errors on resultant data would be to add random noise to the elevations. This form of measurement error is directly quantified as part of the semi-variogram analysis. The nugget effect parameter in the variogram analysis is an estimate of all combined sources of random variation that cannot be explained by spatial variation in the bottom elevation. Strictly speaking, the best estimates of nugget effect are developed from co-located measurements located throughout the area of interest. A close approximate to this situation could be obtained from the crossings of lateral and longitudinal transects, although this would restrict the estimation of this important parameter to a relatively small number of locations, with questionable reliability. As an alternative to this approach, neighboring locations (approximately 10 feet apart) were used as surrogates for co-located samples. The effect of this approach is to overstate the nugget effect, because differences between elevations at pairs of samples separated by 10 feet, usually cross channel, include variation due to both measurement errors and variation in elevation over a 10-foot distance. This was deemed preferable to estimation of pure measurement error from co-located samples because potential understatement of the true nugget effect would tend to inflate the apparent precision of simulation results. It is preferred to overstate uncertainty of erosion estimates than to be distracted with interpretation of small, idiosyncratic features of the data.

Figure 14k shows long- and cross-flow semi-variograms for the 2001 bathymetric surveys. When few sampling locations are available, it is often necessary to make subjective judgments to develop a reasonable model of spatial correlation. For the Lower Passaic River, there are 15 to 20 thousand soundings within each study area, so it was not necessary to make any subjective judgments to fit variogram models. It is particularly interesting to note that in spite of the need to "look" to the nearest samples in order to estimate the nugget effect, the estimated nugget is essentially zero. This essentially zero nugget effect was characteristic of all of the surveys, including the multibeam surveys. This is indicative that concerns over measurement error may be overstated. In general these bathymetric surfaces are very smoothly varying. There is very little evidence of the small-scale variation predicted by ERDC's propagation of errors (Byrnes et al., 2002). If there had been significant measurement error it would have been detected through the variogram analysis.

5.8.2. Random Precision Errors in Recent Bathymetric Measurements

To understand potential uncertainties in measurement precision, point differences in elevation recorded during single beam and multibeam surveys conducted from August 23 to 26, 2007, and September 9 to 22, 2007, respectively, were assessed. This analysis provided an assessment of the uncertainty in the measurement of elevation, assuming that no events occurred in the period between the two surveys. Violation of this assumption would result in overstatement of actual measurement errors. Therefore, these are upper bounds on the measurement errors.

The single beam survey in 2007 contained both transverse tracks (across flow) at 100 feet spacing and three longitudinal tracks (along flow). Three point-based comparisons were made using both the single beam data and the multibeam data. First, the elevations recorded in the transverse and longitudinal single beam tracks were compared to each other at the locations where they intersected. Second and third, the single beam transverse tracks and longitudinal tracks were compared to the multibeam data by determining which single beam depth value was closest to each point of the 3 feet by 3 feet subsampled multibeam data with a maximum allowed distance between the two depths of 1.5 feet. The results of these comparisons are as follows:

- A histogram of depth comparisons at the crossings of along-river and across-river single-beam tracks is given in Figure 23a. The average elevation difference is 0.03 foot (standard deviation = 0.29 foot). About 57 percent of the differences are within ± 0.15 foot.

- A histogram of the comparison of the 2007 single beam transverse tracks to the multibeam data is given in Figure 23b. The average elevation difference is 0.1 foot, the standard deviation is 0.35 foot and the mean absolute error is 0.25 foot.

- A histogram of the comparison of the 2007 single beam longitudinal tracks to the multibeam data is given in Figure 23c. The average elevation difference is 0.03 foot, standard deviation is 0.44 feet, and the mean absolute error is 0.29 feet.

- A map showing the point differences between the longitudinal and transverse tracks single beam compared to the multi beam elevation is shown in Figure 24a through Figure 24m. In this map, the elevation differences are shown in three categories: 1) less than -7 inches (below the 5th percentile), 2) between -7 and 6 inches (5th to 95th percentile) and 3) greater than 6 inches (above the 95th percentile). In general, there is no systematic spatial pattern in the distribution of the differences outside the 5th and 95th percentiles, where differences would typically be expected.

- Observed variances between surveys combine the variances associated with the individual single beam and multibeam surveys. The exact partitioning of the variances among the two surveying methods is unknown. However, the standard deviation of 0.44 foot represents the combined variance of the two surveys and is an upper bound estimate of the variance associated with either survey alone. If it were assumed that each survey was equally precise, one may conclude that the standard deviation for an individual survey is given by $\sqrt{(0.5*(\text{Standard Deviation})^2)}$. This results in standard deviations of 0.2 foot for the point-to-point comparison of cross flow vs. longitudinal transects, 0.25 foot for the cross flow single beam vs. multibeam comparison, and 0.31 foot for the longitudinal single beam vs. multibeam comparison.

These observed errors are somewhat smaller than those typically estimated using error propagation techniques. This is likely the result of the conservatism in some steps of the error propagation methods. These random errors are accounted for by the conditional simulation though the semivariogram analysis and spatial uncertainties and probabilities determined when surveys are compared.

5.8.3. Explanation for Bathymetric Change Analyses' Identification of Erosion at Magnitudes Greater than the Conservative Limits Suggest

The results of linked hydrodynamic and sediment transport modeling in the Lower Passaic River to date has not successfully reproduced the erosional and depositional patterns that have been determined through the comparison of repeat bathymetric surveys. In order to understand why these two approaches to bathymetric change are coming to quite different results, it is important to independently estimate how far the sediment surface can move vertically for flows observed in the Passaic River. This discussion focuses on erosional events. The amount of erosion is estimated by: 1) calculating bed shear stress for several ADCP moorings in the Passaic River collected at different river flows and 2) using measured erosion rates for Passaic River sediments to estimate how much sediment could be removed by those flow events. While the sediments eroded will most likely be deposited in the Lower Passaic River, the location of that deposition may well be different than the location of the sediment erosion. Thus, at least for this analysis, erosion and deposition may be viewed as separate events. Several ADCP deployments have been made in the Passaic River in 2004 and 2005, and ADCP results have recently been posted for a period of low river flow in 2009. ADCP data used in this analysis were collected by USEPA (www.OurPassaic.org), and by R. Chant of Rutgers University on behalf of New Jersey Department of Transportation (NJDOT) (http://marine.rutgers.edu/cool/passaic/). Erosion rates are those reported by Borrowman, *et al.* (2006) and Owens, *et al.* (2006). While ADCP results are available only for about the lower 8 miles of the Lower Passaic River, erosion rates have been estimated for about 14 miles.

Bed shear stress was estimated from Acoustic Doppler Current Profiler (ADCP) data using a drag coefficient based on flow speed measured at a fixed height above the sea bed. Thus

$$\tau 0 = \rho \; CD \; U^2$$

where $\tau 0$ is bed shear stress (Pa), ρ is water density [kilogram per meter cubed (kg/m^3)], CD is a drag coefficient, and U is the flow speed [meter per second (m/s)] measured at about 1 meter (m) off the bed. For this analysis, ρ is estimated as 1,000 kg/m^3 and CD as 0.0017. This value of CD was suggested by R. Chant (Rutgers University, 2010, personal communication) as the best fit value that closed the momentum balance, although he suggested a value in the range 0.0015 to 0.0020 would be reasonable. A higher value of CD of about 0.0031 might be expected for a near-bed flow with a smooth bed, a logarithmic velocity profile and no stratification, so a somewhat smaller value is likely since stratification tends to reduce the bed shear stress for a given flow. The bottom-mounted ADCP instruments report flow speeds at various elevations above the bed, and for this analysis the flow speed reported at an elevation of about 0.8 to 1.0 m was used, which is bin number 2 or 3 depending on the instrument used. The ADCP data used here were inspected to remove obvious instrumental errors, but no other processing of speed data was done.

Given the generally small number of moorings available and their limited time duration, not all possible combinations of tide and river discharge are sampled in all parts of the river. However, ADCP records exist for the high discharge event of 11,700 cubic feet per second (cfs) in early April 2005 (file 02BY1000.000 at mooring buoy 1), for intermediate discharge events of 4,110 and 5,470 cfs in

December 2004 and January 2005 respectively (file PASS2001 at DOT2 mooring M2b), and for lower flow events of 3,690 and 3,940 cfs in September 2004 (file RDI2484_dot1 at DOT1 mooring M2).

File 02BY1000.000 extended from February 2 to May 5, 2005 and included a discharge event of 11,700 cfs on April 5, 2005 (Figure 25a). The unit reports a flow speed every 15 minutes and the reported flow speeds in the 3rd bin, or at 0.8 m above the bottom, were used in this analysis. The maximum flow speed at 0.8 m above the bottom for this time was 1.1 m/sec on April 5, 2005 during the discharge event with flow speeds over 1 m/sec occurring for a total of 16 hours during this discharge event. Maximum bed shear stress during this flow event was about 2 Pa, bed shear stress was above 1.6 Pa for nearly 24 hours over five tidal cycles, bed shear stress was above 0.8 Pa for about 155 hours over about 23 tidal cycles and bed shear stress was above 0.4 Pa for about 280 hours over about 28 tidal cycles

File PASS2001 extended from November 19, 2004 to January 24, 2005 and included a discharge event of 4,110 cfs on December 2, 2004 and a discharge event of 5,470 cfs on January 16, 2005 (Figure 25b). The unit reports a flow speed every 30 minutes and the reported flow speeds in the 2nd bin, or at 0.9 m above the bottom, were used in this analysis. The maximum flow speeds were over 0.80 m/sec at 0.9 m above the bottom, but for this record flow speed was not directly related to river flux. Maximum bed shear stresses over 1.0 Pa were determined for brief periods of five tidal cycles. Bed shear stresses were over 0.8 Pa for 29 hours during this record, and they were over 0.4 Pa for 263 hours. The three other long ADCP records that exist for parts of this time period give somewhat similar results.

File RDI2484_dot1 extended from August 18 to October 21, 2004 and included discharge events of 3,690 cfs on September 20 and of 3,940 cfs on September 30, 2004 (Figure 25c). Two other smaller events of about 2,000 cfs occurred in the first half of the record. The unit reports a flow speed every 20 minutes and the reported flow speeds in the 3rd bin, or at 0.8 m above the bottom, were used in this analysis. The maximum flow speed for this record was about 0.97 m/sec during one tidal cycle, and three tidal cycles have maximum flows over 0.7 m/sec. The maximum calculated bed shear was 1.6 Pa and bed shear stress over 0.8 Pa for 7 hours during this record, and they were over 0.4 Pa for over 101 hours. One other ADCP record exists for a portion of this time period and that record gives somewhat similar results.

In summary, the ADCP results suggest that bed shear stresses can be over 1.6 Pa during high-discharge events, and that bed shear stresses decrease as the discharge decreases. However, bed shear stress can be over 0.8 Pa for significant time periods during these events, with the amount of time at this bed shear stress decreasing as river flow decreases. Bed shear stress can be over 0.8 Pa for 7 hours during flow events of about 4,000 cfs.

We can estimate depth of sediment erosion by flow by using sediment properties determined through the Sedflume analysis which determined erosion rate in centimeter per second (cm/sec) as a function of bed shear stress in Pa (Borrowman, *et al.*, 2006). A total of 28 cores were analyzed at 14 different sites. Coring sites were chosen to characterize different sediment types, and some of the cores are in the general vicinity of the ADCP moorings. However, for the purposes of this memo the erosion properties of the sediments are described in general rather than trying to associate any particular core with an

ADCP record. The cores range in length from about 30 to over 45 centimeter (cm). For each core there is a plot in Borrowman, *et al.* (2006) that shows the results of measuring erosion rates at different bed shear stresses. Again, for this memo the primary interest is in sediments that will erode at a rate greater than 1×10^{-3} cm/sec, since sediment eroding at this rate for 7 hours (2.52×10^{4} seconds) will have an elevation change of 25.2 cm (about 0.8 foot).

For many of the cores studied by Borrowman, *et al.* (2006), erosion rates at 0.8 Pa are greater than about 1×10^{-3} cm/sec for 33 cm to over 40 cm (about 1.1 foot to 1.3 foot; *e.g.*, cores P09B, P13A and P15B). In some cores, for example cores P02A, P02B, P03A, P04A, P04B, P07A, P09A, P11A and P11B, there are fairly thin layers with lower erosion rates in the upper parts of the cores, but higher erosion rates again occur to depths greater than 25 cm (0.8 foot) beneath the resistant layers.

While this is not a extensive analysis of the erosional properties of Lower Passaic River sediments or of flow characteristics in the Lower Passaic River, this review is sufficient to indicate that bed shear stresses of over 0.8 Pa for 7 hours (the magnitude and duration of bed shear stress calculated from August to October 2004) should be able to erode many Lower Passaic River sediments to depths of 0.8 foot or more during discharge events of about 4,000 cfs or greater. This magnitude of depth change is generally consistent with those observed through the comparison of repeat bathymetric surveys, and there were over 130 days with river discharges over 4,000 cfs between 1989 and 2004, the time interval used in the bathymetric survey analysis.

Given that this simple estimate of bed shear stress and sediment erosion supports the bathymetric change analysis, the question then becomes why the hydrodynamic / sediment transport model is calculating much smaller bathymetry changes than the bathymetric change analysis. If it is assumed that the hydrodynamic model is correctly calculating bed shear stress, the assumptions being used for sediment erosion need to be revisited. These steps include understanding the quality of discrete erosion rate measurements (*e.g.*, possible sediment compaction during core collection and analysis or sediment changes due to consolidation or other effects such as loss of gas bubbles following core recovery), understanding how sediments are consolidated in the model, and understanding the role of sediment variability in calculating the distribution of sediment erosion properties on the sea bed. For example, averaging together many core profiles may effectively increase the apparent sediment resistance to erosion while single, weak layers such as those related to leaves or gas in the sediment, might, in fact, control the location and amount of sediment erosion.

5.9. Validation of Conditional Simulation and Establishment on Probabilities of Bathymetric Change Thresholds

Validation assessment of bathymetric interpolation by conditional simulation was done by:

- Creating a 1990's- type single beam survey by sampling the 2007 mutibeam data along the 1995 bathymetry survey transect locations. This was referred to as "2007 at 1995 transect locations".

- Performing conditional simulation of the 2007 at 1995 transect locations, on a 6-foot by 9-foot grid.

- Comparing the conditional simulation results with the true 2007 multibeam data at the unsampled, simulated grid points.

The validation of the conditional simulation method to spatial interpolation of single beam bathymetry data was conducted to evaluate the performance of the method in determining the elevation and associated probability especially in locations unaligned with the single beam transect tracks. The conditional simulation approach was applied to the 2007 on 1995 transect locations to simulate elevations at grid locations where actual elevation measurements exist from the multibeam survey. A histogram of the differences between the average simulated surface and the actual surface is given in Figure 26a. The mean absolute difference is 0.27 feet, and the 5th and 95th percentile of the differences are between -6 and 6 inches, respectively, a result that is comparable to that obtained in the single beam versus multibeam point comparison. Figure 26b compares the histogram of the validation difference for the conditional simulation from Figure 26a to the difference from the point-based single beam and multibeam comparison from Figure 26b. The distributions of the differences are comparable, and show more than 65 percent of the differences are within 3 inches or less.

The degree of agreement in the validation analysis was further tested to determine the probability threshold for a 5 percent false positive rate for various elevation differences. A false positive (or Type I) error is defined in this context as simulating a substantial erosion change beyond a depth difference cutoff when in truth such a difference shouldn't occur. Figure 27 presents a summary of the results, highlighting in yellow the probability thresholds for a 5 percent false positive rate. For a 6-inch erosion cutoff, a 5 percent false positive rate occurs when the probability threshold is set at approximately 50 percent. Therefore, when evaluating the results of the bathymetric change analysis performed using conditional simulation, a significant bathymetric change of 6 inches should be assessed at the 50 percent probability threshold. Note that the probability threshold for a 5 percent false positive rate increases as the thickness criterion decreases. Figure 28a through 28e shows the spatial locations where the false positive results occurred for a 6-inch difference cutoff with at least 50 percent probability. The false positive locations are preferentially located close to bridges where uncertainties in bathymetry data are expected to be higher due to steep surfaces and positioning signal interference caused by the bridge superstructure. Errors are also observed at the edges of the simulation where data coverage and interpolation are less robust. Overall, these observations strongly suggest that for 6 inches and more of elevation change, the 70 percent probability threshold used to interpret bathymetric change maps is less protective of the river bed.

6.0. Preliminary Comparison of 2007 and 2008 Multibeam Surveys

Roger Flood of Stony Brook University performed a preliminary comparison of the 2007 and 2008 multibeam surveys in five segments (A to E) along the Passaic River (Figure 29) in order to determine if there were patterns in bathymetric change that could be discerned based on these high-resolution data sets. This new data set can help to understand the nature and origin of bathymetric changes estimated from the single beam surveys. The results of the comparison are as follows:

- For segment A around RM13.6, sun-illuminated bathymetry images show sand waves with some mounds in 2007 (Figure 30a), but a smoother bed with no sand waves, although with many more mounds visible, in 2008 (Figure 30b). Bathymetric comparisons along the longitudinal transect in segment A (drawn in blue in Figures 30a and 30b) show the 2008 surface to be deeper with no sand waves relative to 2007 (Figure 30c). This segment was eroded between the 2007 and 2008 survey periods; Figure 30d shows the extent of erosion, including the erosion of the sand waves.

- For river segment B around RM9.75, sun-illuminated bathymetry images in 2007 (Figure 31a) and 2008 (Figure 31b) show a coarse bed beneath the bridge in the center of the segment. However, the coarse bed in the 2008 bathymetric surface in this area is about 0.2 feet deeper (Figure 31c) than in 2007. Figure 31d shows the extent of the scour and deposition around the bridge area based on the difference between 2008 and 2007 survey results. This uncertainty around bridges observed between the multibeam surveys is consistent with that reported above in the conditional simulation validation analysis. River segment C around RM8.85 (Figure 32a though d) further illustrate the effect of bridges and other structures on bathymetry uncertainty. Figure 32d also shows a sand wave field (shown in red) south of the bridge that developed between 2007 and 2008.These results also suggest a difference of 0.1 to 0.2 foot in the area around the bridge structure. Assuming that the absolute elevation of these areas did not change between surveys, these observations suggest a survey-to-survey absolute error of 0.2 feet or less. As can be seen in the survey comparisons reported in Chapter 11 of the Conceptual Site Model, this absolute error is small in comparison to many of the observed elevation changes.

- For river segment D around RM4.05, the 2007 sun-illuminated images show large bed scours (Figure 33a), but these areas are filled in 2008 (Figure 33b), an illustration of the sediment dynamics in the river. A transverse transect in this segment shows both erosion of more than 6 inches and deposition of more than 1 foot moving from the Northwest bank to the Southeast bank of the river (Figure 33c). Significant erosion and deposition changes can be observed throughout this segment as shown in Figure 33d.

- For river segment E around RM1.6, the sun-illuminated images in 2007 show a smooth surface with some local depressions (Figure 34a). In 2008, the images show an irregular surface with some local depressions (Figure 34b). A transect along the flow direction shows small deposition at one end and large erosion change of more than 2 feet at the other (Figure 34c). Note that this elevation change is more than an order-of-magnitude greater than the offset between surveys suggested above. Overall this area is erosional with some patches of significant deposition along the edges. Part of the erosion observed can be attributed to ship activities.

This preliminary analysis identifies areas of significant bathymetric change between the two multibeam surveys, despite their relative closeness in time (approximately 1 year). In many areas, particularly at RM4.05 and RM1.6 (Areas D and E), the scale of change was six inches or more. Additionally, the results show geomorphic changes that correspond with the elevation changes.

The geomorphic changes apparent in the multibeam survey can be considered characteristic of the processes that cause the elevation changes noted in the historical bathymetric surveys. That is, the same processes that cause these features to come and go also cause the changes in absolute elevation. However, the spacing of the bathymetric survey cross-sections (100 feet apart) and the scales of the geomorphic features (much less than 100 feet) are very different. Thus, it is not the appearance of these features that is causing the observations of bathymetric change. Additionally, changes of statistical significance generally require multiple bathymetric nodes to correlate; thus the scale of the detected bathymetric changes between surveys generally spans more than several hundred feet.

With that said, the occurrence of wave features and the appearance and disappearance of large sediment mounds can be considered part of an active layer of sediment that is mobile and is subject to routine movements. This layer is part of the active volume of sediments into which solids delivered from the head-of-tide, solids delivered from Newark Bay, and solids eroded from deeper layers are mixed. It is this layer that is still highly contaminated with dioxins, PCBs, metals, PAHs, pesticides and other contaminants of concern.

7.0. Summary

This attachment is intended to document the simulation procedure that was used to support analyses reported in Chapter 11 of the CSM. The methods are relatively new to the environmental sciences, but are well documented in theoretical and practical developments in the geological engineering and geostatisitics literature. The advantages of the conditional simulation approach are worthwhile when data are correlated and inferences are needed at scales other than the original data. Essentially, any complicated non-linear transfer functions of the surface can be accommodated. In this instance, separate estimates of volume of material eroded (deposited) were of interest and interpolated maps did not provide adequate understanding of the magnitude and frequency of erosion at sub-model grid scales. These methods provided an improved understanding of these processes.

It has been asserted that, because year over year erosion and deposition patterns are more dynamic than predicted by sediment transport models, the results may not be believable. This assertion is difficult to evaluate directly because deterministic models tend to identify circumstances of net deposition or net erosion over relatively large areas. However, the results of the conditional simulation indicate that within areas that are net depositional, there are often both erosional and depositional events, suggesting a continued mixing of a relatively thick (typically more than 6 inches and often more than a foot) layer of sediment. It is not clear that these results are necessarily inconsistent. In efforts to develop a qualitative assessment of the degree to which the conditional simulation agrees with more conventional wisdom, it is interesting to consider the net erosion and deposition estimated by the conditional simulation over the longer period of time from 1989 through 2007.

Figure 35 shows net erosion and deposition of sediments in the Lower Passaic based on the conditional simulation. The color scales represent the magnitude of erosion and deposition, and the strength of the colors represent the probability that these magnitudes of erosion occurred.

Dark red areas represent areas where it is 95 percent likely that greater than 1 foot of erosion occurred and dark blue areas indicate areas where it is 95 percent likely that more than a foot of deposition occurred between 1989 and 2007. It can be seen that this analysis places these high magnitude areas along outside bends in the higher energy sections of the river. Similarly, areas with thick accumulations of sediments at high levels of confidence are in the lower 2 miles of the river where energies are lower, as well as along inside bends and within the main channel areas adjacent to outside bends with steep eroding slopes. Areas with 6 inches to 1 foot of erosion with 95 percent level of probability are also located in geomorphologically reasonable areas at the heads and tails of strongly erosional outer bends. Erosional and depositional areas in the upper portion of the river tend to be smaller and more variable in size with mixtures of erosion and deposition apparently influenced by locations of bridge abutments and as well as less extreme inside and outside bends.

These results identify erosional and depositonal features that are qualitatively in the “right” places. Further they do not show any indication of artificial banding or discontinuities that would be expected if the data contained the biases and large systematic errors with which the ERDC reviewer is concerned (USACE ERDC, 2010). Importantly, these analyses also provide quantitative bounds on the precision of estimated erosion and deposition, providing managers with an understanding of the magnitude of erosion and deposition that has been observed and the strength of evidence supporting observed dynamics. The richness of this analysis is unavailable through other means.

References

Barabás, N., Goovaerts, P., and Adriaens,P. 2001. Geostatistical Assessment and Validation of Uncertainty for Three-Dimensional Dioxin Data from Sediments in an Estuarine River. Environ. Sci. Technol., 2001, 35 (16), pp 3294–3301.

Barabas, N., J.R. Wolfe and J.W. Kern. 2009. The influence of uncertainty in models of dredged sediment volumes vs. sampling density. Fifth International Conference on Remediation of Contaminated Sediments. February 2-5, 2009 Jacksonville, Florida.

Borgman, L.E., M. Taheri, and R. Hagan. 1984. Three-dimensional frequency-domain simulations of geological variables. In G. Verly, M. David, A. G. Journel, and A. Marechal, Editors, *Geostatistics for Natural Resources Characterization,* volume 2, pages 517-541.

Borrowman, T.D., Smith, E.R., Gailani, J.Z., and Caviness, L., 2006. Erodibility Study of Passaic River Sediments Using USACE Sedflume. USACE ERDC TR-06-7.

Byrnes, M.R., Baker, J.L, and Li, F. 2002. Quantifying potential measurement errors and uncertainties associated with bathymetric change analysis. Coastal and Hydraulic Engineering Technical Note CHETN-IV-50. U.S. Army Engineer Research and Development Center, Vickburg, MS. (http://chl.wes.army.mil/library/publications/chetn).

Cressie, N.A.C. 1993. Statistics for Spatial Data, Revised Edition. John Wiley and Sons. 928 pp.

Deutch, C. V. and A.G. Journel. 1998. *GSLIB: Geostatistical Software Library and Users Guide.* Oxford University Press, New York, 370 p.

Driscoll, T. A., A MATLAB toolbox for Schwarz–Christoffel mapping, ACM *Trans. Math. Soft*., 22 (1996), pp. 168–186.

GBA 2010. GBA's Response to USEPA's bathymetry survey questions. Received 3/23/2010.

Goovaerts, P.M. 1997. *Geostatistics for Natural Resources Evaluation.* Applied Geostatistics Series, Oxford University Press, New York. 483p.

Isaaks, E. H. and Srivastava R. M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, USA. 592 pages.

Journel, A.G. and Ch. J. Huijbregts. 1978. *Mining Geostatistics.* Academic Press, New York. 600 p.

Kern, J.W. and L.E. Borgman. 1997. Frequency domain conditional simulation using the fast Fourier transform and recently developed consistency conditions. United States Department of Interior, Office of Surface Mining Reclamation and Enforcement, Technology Transfer. http://www.techtransfer.osmre.gov/NTTMainSite/Library/pub/fdcs.pdf. (Last visited 4/19/2009).

Kern, J.W. 2005. Relationship between dredge area footprint and probability of excedence for Tri+ PCBs 10 mg/kg action limits. March 9, 2005 presentation to the Acting Regional Deputy administrator, USEPA Region 2.

Kern, J.W., J. Wolfe and N. Barabas. 2008. Evaluation of increased sampling density for refinement of 30 percent dredge prism design in Upper OU-3. Fox River PCBs Superfund Site. Technical working group memorandum. 28 pp.

Kern, J.W. J.R. Wolfe and N. Barabas. 2009. Sampling Density for Refinement of Dredge Prism Design in the Lower Fox River, Wisconsin OU 2-5. Fifth International Conference on Remediation of Contaminated Sediments. February 2-5, 2009 Jacksonville, Florida.

OSI 2010. OSI's Response to USEPA's bathymetry survey questions. Received 3/26/2010.

Owens, M., Cornwell, J.C., Suttles, S.E. and Dickhudt, P., 2006. Passaic River Erosion Testing and Core Collection: Field Report and Data Summary. Chesapeake Biogeochemical Associates. Subcontract No. KC-ACE2002-31

Ripley, B.D. 1987. *Stochastic Simulation.* John Wiley and Sons, New York, NY.

Saito, H. and P. Goovaerts, 2002, Accounting for Measurement Error in Uncertainty Modeling and Decision Making using Indicator Kriging and p-field Simulation: Application to a Dioxin Contaminated Site. Environmetrics, 13(5-6): 555-567.

Srivastava, R.M. 1992. Reserve characterization with probability field simulation. In SPE *Annual conference and Exhibition, Washington, D.C.,* Number 24753, pages 927-938, Washington D.C., Society of Petroleum Engineers.

Srivastava, R.M. and R. Froidevaux. 2005. Probability field simulation: A retrospective. In: Leuangthong, O. and C.V. Deutsch (eds). *Geostatistics Banff 2004.* 55-64. Springer. Netherlands 2005.

The Math Works. 1998. MATLAB, The Language of Technical Computing, Version 5.2. www.themathworks.com. Last visited June 19, 2009.

USACE ERDC, 2010. Draft memorandum for record: Survey Accuracies on the Passaic River. Submitted to USEPA on 3/2/2010.

USEPA. 1989. Methods for evaluating the attainment of cleanup standards. Volume 1: Soils and Solid Media. United States Environmental Protection Agency. Office of Policy, Planning and Evaluation. Washington, DC 20460. EPA230/02-89-042.

Table 1. Potential errors associated with bathymetric surveys and classification into their potential influence on measurements.

Potential Error Source	Error Type
Horizontal Accuracy	Reproducibility
Vertical Accuracy	Reproducibility
Vertical Benchmarks	Global Bias (vertical offset)
Heave, Pitch, Roll Compensation	Reproducibility
Bar Checks	Bias (vertical offset)
Sound Velocity Casts	Local Bias (vertical offset)
Vertical Datum	Global Bias (vertical offset)
Latency/Patch Test	Local Bias (vertical offset)
Squat/ Settlement Test	Local Bias (vertical offset)